文章编号:1671-9352(2007)06-0065-04

四元数矩阵的 QR 分解及等式约束最小二乘问题

赵建立[1,2],李 莹[2],张丽梅[2]

(1. 华东师范大学 数学系,上海 200062; 2. 聊城大学 数学科学学院,山东 聊城 252059)

摘要:利用 Givens′变换给出了四元数矩阵的 QR 分解,并利用复表示和 QR 分解解决了 2-范数下的四元数矩阵的等式约束最小二乘问题.

关键词:四元数矩阵;最小二乘;复表示;分解

中图分类号:O151.21 **文献标识码**:A

QR decomposition and the equality constrained least squares problem over a quaternion field

ZHAO Jian-li[1,2], LI Ying[2] and ZHANG Li-mei[2]

(1. Department of Math., EastChina Normal University, Shanghai 200062, China;
2. School of Mathematics Science, Liaocheng Univ., Liaocheng 252059, Shandong, China)

Abstract: By using transformation, the decomposition of quaternion matrices is presented and the equality constrained least squares problem of quaternion matrices with complex presentation and decomposition is studied.

Key words: quaternion matrices; equality constrained least squares problem; complex presentation of quaternion; decomposition

如同复数域上的矩阵一样,最小二乘问题在四元数矩阵理论与应用中占有十分重要的地位. [1]利用复表示研究了四元数矩阵的最小二乘问题

$$\| \boldsymbol{A}x - b \|_2 = \min,$$

给出了最小二乘解及最小范数解的表达式. [2]针对 Frobenius 范数,利用复表示和四元数矩阵的 GSVD 方法研究了四元数矩阵的等式约束最小二乘(ECLS)问题

$$\| \boldsymbol{A}x - b \| = \min \quad \text{s.t.} \quad \boldsymbol{B}x = d, \tag{1}$$

给出了解存在的充要条件及解的表达式.本文利用[3]中介绍的四元数矩阵的 Givens′变换给出了四元数矩阵的 QR 分解,并利用复表示、QR 分解及[4]中的某些结果解决了 2-范数下的四元数矩阵的 ECLS 问题.本文中 $\boldsymbol{A}\in Q^{m\times n}$, $\boldsymbol{B}\in Q^{p\times n}$, $b\in Q^m$, $d\in Q^p$, $\mathrm{rank}(\boldsymbol{A})=n$, $\mathrm{rank}(\boldsymbol{B})=p$.

1 预备知识

令 $\mathbf{R}$ 表示实数域,$\mathbf{C}$ 表示复数域,Q 表示四元数体. $Q^{m\times n}$ 表示四元数体 Q 上全体 $m\times n$ 矩阵的集合,$U^{n\times n}$ 表示全体 $n\times n$ 四元数酉矩阵的集合.对于 $\boldsymbol{A}\in Q^{m\times n}$,$\boldsymbol{A}^{\mathrm{H}}$ 表示 $\boldsymbol{A}$ 的共轭转置,$\boldsymbol{A}^{+}$ 表示 $\boldsymbol{A}$ 的广义逆矩阵.

对任意 $x = x_0 + x_1 i + x_2 j + x_3 k\in Q$,

定义 $x^{\delta} = \begin{bmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} \in \mathbf{C}^{2\times 2}$ 为 x 的复表示.

收稿日期:2006-11-02
基金项目:山东省自然科学基金资助项目(2004ZX13)
作者简介:赵建立(1964-),男,教授,博士研究生,研究方向:数值代数. E-mail: zhaojl1964@gmail.com

对于 $\boldsymbol{A}=(a_{ij})\in Q^{m\times n}$,定义 $\boldsymbol{A}^{\delta}=((a_{ij})^{\delta})\in \mathbf{C}^{2m\times 2n}$为$\boldsymbol{A}$ 的复表示阵.

定理 1.1[1,6] (1) 若 $\boldsymbol{A},\boldsymbol{B}\in Q^{m\times n}$, $a\in R$,则$(\boldsymbol{A}+\boldsymbol{B})^{\delta}=\boldsymbol{A}^{\delta}+\boldsymbol{B}^{\delta}$, $(a\boldsymbol{A})^{\delta}=a\boldsymbol{A}^{\delta}$,

(2) 若 $\boldsymbol{A}\in Q^{m\times n}$, $\boldsymbol{B}\in Q^{n\times s}$,则$(\boldsymbol{AB})^{\delta}=\boldsymbol{A}^{\delta}\boldsymbol{B}^{\delta}$,

(3) 若 $\boldsymbol{A}\in Q^{m\times n}$,则$(\boldsymbol{A}^{\mathrm{H}})^{\delta}=(\boldsymbol{A}^{\delta})^{\mathrm{H}}$, $(\boldsymbol{A}^{+})^{\delta}=(\boldsymbol{A}^{\delta})^{+}$,

(4) 若 $\boldsymbol{A}\in Q^{n\times n}$,则 $\boldsymbol{A}$ 可逆当且仅当$\boldsymbol{A}^{\delta}$ 可逆且$(\boldsymbol{A}^{\delta})^{-1}=(\boldsymbol{A}^{-1})^{\delta}$,

(5) $\boldsymbol{A}\in U^{n\times n}$的充要条件是$\boldsymbol{A}^{\delta}$ 为酉阵.

定义 1.1[6] 令 $\boldsymbol{A}\in Q^{m\times n}$,四元数矩阵 $\boldsymbol{A}$ 的秩 $\mathrm{rank}(\boldsymbol{A})$定义为

$$\mathrm{rank}(\boldsymbol{A})=\frac{1}{2}\mathrm{rank}(\boldsymbol{A}^{\delta}).$$

$Q_r^{m\times n}$ 表示全体秩为 r 的 $m\times n$ 四元数矩阵的集合.

定义 1.2[5] 若函数 $v: Q^n\to \mathbf{R}^+$ 满足对任意 $x,y\in Q^n$, $\alpha\in\mathbf{R}$,有

(1) $x\neq 0\Rightarrow v(x)>0$, (2) $v(\alpha x)=|\alpha|v(x)$, (3) $v(x+y)\leq v(x)+v(y)$,

则称 v 为Q^n 上的范数.

如对于 $x=(x_1,x_2,\cdots x_n)^{\mathrm{T}}\in Q^n$, $\|x\|_F=\sqrt{\sum\limits_{j=1}^{n}|x_j|^2}$,为向量 $\boldsymbol{x}$ 的 Frobenius 范数.

定义 1.3[1] 一个从 $Q^{m\times n}$到 $\mathbf{R}^+$ 的映射 v 叫做$Q^{m\times n}$上的四元数矩阵范数,如果它满足

(1) $\boldsymbol{A}\neq 0\Rightarrow v(\boldsymbol{A})>0$, (2) $v(\alpha\boldsymbol{A})=|\alpha|v(\boldsymbol{A})$, (3) $v(\boldsymbol{A}+\boldsymbol{B})\leq v(\boldsymbol{A})+v(\boldsymbol{B})$.

其中 $\boldsymbol{A},\boldsymbol{B}$ 为 $Q^{m\times n}$中任意的矩阵,α 为任意复数.

对于 $\boldsymbol{A}=(a_{ij})\in Q^{m\times n}$,令 $\|\boldsymbol{A}\|_F=\sqrt{\sum\limits_{i,j}|a_{ij}|^2}$, $\|\boldsymbol{A}\|_2=\sqrt{\lambda_{\max}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}=\delta_{\max}(\boldsymbol{A})$,其中 $\lambda_{\max}$与 $\delta_{\max}$分别表示最大特征值与最大奇异值.易证 $\|\boldsymbol{A}\|_F$, $\|\boldsymbol{A}\|_2$ 均为四元数矩阵范数,且 $\|\boldsymbol{A}\|_F=\|\boldsymbol{A}^{\delta}\|_F$, $\|\boldsymbol{A}\|_2=\|\boldsymbol{A}^{\delta}\|_2$.称 $\|\cdot\|_F$ 与 $\|\cdot\|_2$ 分别为四元数矩阵的 Frobenius 范数与2-范数.

2 四元数矩阵 QR 分解

定理 2.1[3] 对于非零向量 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}\in Q^2$, 令 $\boldsymbol{G}=\begin{bmatrix}\bar{c} & s\\ -\bar{s} & c\end{bmatrix}\in Q^{2\times 2}$,其中 $s=-\delta\dfrac{\bar{x}_2}{\|x\|_F}$, $c=\delta\dfrac{\bar{x}_1}{\|x\|_F}$, $|\delta|=1$,并且当 x_1, x_2 在实数域上线性相关时,δ 为任意四元数,否则,δ 在集合 $\sum=\left\{\delta\in Q\,\middle|\,\delta=\dfrac{\alpha x_1+\beta x_2}{|\alpha x_1+\beta x_2|},\alpha,\beta\in R,\ |\alpha|+|\beta|>0\right\}$中选取.则 $\boldsymbol{G}$ 是酉阵且$\boldsymbol{G}^{\mathrm{H}}x=u=\delta\|x\|_F(1,0)^T$.

定理 2.2 设 $\boldsymbol{A}\in Q^{m\times n}$且 $\mathrm{rank}(\boldsymbol{A})>0$,则存在 m 阶酉阵$\boldsymbol{Q}$ 和上三角矩阵$\boldsymbol{R}$,使得

$$\boldsymbol{QA}=\begin{bmatrix}\boldsymbol{R}\\ 0\end{bmatrix}.$$

证明 利用定理2.1的 Givens′变换给出定理的一个构造性证明.

对矩阵的列数作归纳法.当 $n=1$ 时,设 $\boldsymbol{A}=(a_{11},a_{21},\cdots a_{m1})^{\mathrm{T}}$,

令 $x_1=(a_{m-1,1},a_{m1})^{\mathrm{T}}$,且令 $\boldsymbol{Q}_1=\begin{bmatrix}1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \bar{c}_1 & s_1\\ & & & -\bar{s}_1 & c_1\end{bmatrix}$, 则 $\boldsymbol{Q}_1\boldsymbol{A}=[a_{11},\cdots,a_{n-2,1},\delta_1\|x_1\|_F\quad 0]^{\mathrm{T}}$.

令 $x_2=(a_{m-2,1},\delta_1\|x_1\|_F)^{\mathrm{T}}$,且令 $\boldsymbol{Q}_2=\begin{bmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \bar{c}_2 & s_2 & \\ & & & -\bar{s}_2 & c_2 & \\ & & & & & 1\end{bmatrix}$, 则 $\boldsymbol{Q}_2\begin{bmatrix}a_{11}\\ \cdots\\ a_{m-2,1}\\ \delta_1\|x_1\|_F\\ 0\end{bmatrix}=\begin{bmatrix}a_{11}\\ \cdots\\ \delta_2\|x_2\|_F\\ 0\\ 0\end{bmatrix}$.

令 $x_3=(a_{m-3,1}\ \delta_2 \| x_2 \|_F)^{\mathrm T}$,且令

$$\boldsymbol{Q}_3=\begin{bmatrix}1 &&&&&& \\ &\ddots&&&&& \\ &&1&&&& \\ &&&\bar c_3 & s_3 && \\ &&&-\bar s_3 & c_3 && \\ &&&&&1& \\ &&&&&&1\end{bmatrix},$$

则 $\boldsymbol{Q}_3\begin{bmatrix}a_{11}\\ \cdots \\ \delta_2 \| x_2 \|_F \\ 0 \\ 0\end{bmatrix}=\begin{bmatrix}a_{11}\\ \cdots \\ \delta_3 \| x_3 \|_F \\ 0 \\ 0 \\ 0\end{bmatrix}$.

如此继续,经过 $m-1$ 次 Givens′变换可得$[\delta_{m-1} \| x_{m-1} \|_F,\ 0,\cdots 0]^{\mathrm T}$.

令 $\boldsymbol{R}=\delta_{m-1} \| x_{m-1} \|_F$,$\boldsymbol{Q}=\boldsymbol{Q}_{m-1}\cdots\boldsymbol{Q}_2\boldsymbol{Q}_1$,显然 $\boldsymbol{Q}$ 为 m 阶酉阵,且 $\boldsymbol{QA}=\begin{bmatrix}\boldsymbol{R}\\ 0\end{bmatrix}$,即 $n=1$ 时成立.假设结论对于列数为 $n-1$ 的矩阵成立,对于 $\boldsymbol{A}=(a_{ij})\in Q^{m\times n}$,由 $n=1$ 证明知存在 m 阶酉阵 $\boldsymbol{Q}_1$ 使得 $\boldsymbol{Q}_1\boldsymbol{A}=\begin{bmatrix}a_{11}^{(1)} & y^{\mathrm T}\\ 0 & \boldsymbol{A}_1\end{bmatrix}$,其中 $\boldsymbol{A}_1$ 为 $\boldsymbol{Q}_1\boldsymbol{A}$ 的$(m-1)\times(n-1)$子阵,由归纳假设存在 $m-1$ 阶酉阵 $\widetilde{\boldsymbol{Q}}$ 和上三角矩阵$\widetilde{\boldsymbol{R}}$,使得 $\widetilde{\boldsymbol{Q}}\boldsymbol{A}_1=\begin{bmatrix}\widetilde{\boldsymbol{R}}\\ 0\end{bmatrix}$.令 $\boldsymbol{Q}_2=\begin{bmatrix}1 & \\ & \widetilde{\boldsymbol{Q}}\end{bmatrix}$,则 $\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A}=\begin{bmatrix}a_{11}^{(1)} & y^{\mathrm T}\\ 0 & \widetilde{\boldsymbol{R}}\\ 0 & 0\end{bmatrix}$,令 $\boldsymbol{Q}=\boldsymbol{Q}_2\boldsymbol{Q}_1$,显然 $\boldsymbol{Q}$ 为 m 阶酉阵.令 $\boldsymbol{R}=\begin{bmatrix}a_{11}^{(1)} & y^{\mathrm T}\\ 0 & \widetilde{R}\end{bmatrix}$,$\boldsymbol{R}$ 为上三角阵.所以结论对列数为的矩阵也成立.由归纳法知该结论成立.

3　四元数矩阵的 ECLS 问题

对于(1),我们考虑其复表示阵的等式约束最小二乘

$$\| \boldsymbol{A}^{\delta}y-b^{\delta} \|_2=\min \quad \text{s.t.} \quad \boldsymbol{B}^{\delta}y=d^{\delta}. \tag{2}$$

定理 3.1　四元数矩阵的等式约束最小二乘问题(1)有一解 x 当且仅当其复表示阵的等式约束最小二乘问题(2)有一解 $y=x^{\delta}$.

证明　设(2)有解 $y=x^{\delta}$,对 $\boldsymbol{B}^{\mathrm H}$ 进行 QR 分解

$$\boldsymbol{Q}^{\mathrm H}\boldsymbol{B}^{\mathrm H}=\begin{bmatrix}\boldsymbol{R}\\ 0\end{bmatrix}\begin{matrix}p\\ n-p\end{matrix},$$

则
$$(\boldsymbol{Q}^{\delta})^{\mathrm H}(\boldsymbol{B}^{\delta})^{\mathrm H}=\begin{bmatrix}\boldsymbol{R}^{\delta}\\ 0\end{bmatrix}\begin{matrix}2p\\ 2n-2p\end{matrix}. \tag{3}$$

令
$$\boldsymbol{A}^{\delta}\boldsymbol{Q}^{\delta}=[\boldsymbol{A}_1^{\delta},\boldsymbol{A}_2^{\delta}],\ (\boldsymbol{Q}^{\delta})^{\mathrm H}y=(\boldsymbol{Q}^{\mathrm H})^{\delta}x^{\delta}=\begin{bmatrix}t^{\delta}\\ z^{\delta}\end{bmatrix}\begin{matrix}2p\\ 2n-2p\end{matrix},$$

则
$$\boldsymbol{A}^{\delta}y=\boldsymbol{A}^{\delta}\boldsymbol{Q}^{\delta}(\boldsymbol{Q}^{\delta})^{\mathrm H}y=[\boldsymbol{A}_1^{\delta},\boldsymbol{A}_2^{\delta}]\begin{bmatrix}t^{\delta}\\ z^{\delta}\end{bmatrix}=\boldsymbol{A}_1^{\delta}t^{\delta}+\boldsymbol{A}_2^{\delta}z^{\delta}. \tag{4}$$

由(3)得
$$((\boldsymbol{Q}^{\delta})^{\mathrm H}(\boldsymbol{B}^{\delta})^{\mathrm H})^{\mathrm H}=\begin{bmatrix}\boldsymbol{R}^{\delta}\\ 0\end{bmatrix}^{\mathrm H},$$

即
$$\boldsymbol{B}^{\delta}\boldsymbol{Q}^{\delta}=[(\boldsymbol{R}^{\delta})^{\mathrm H},0],$$

所以 $$\boldsymbol{B}^{\delta}=[(\boldsymbol{R}^{\delta})^{\mathrm{H}},0](\boldsymbol{Q}^{\delta})^{\mathrm{H}}.$$

$$\boldsymbol{B}^{\delta}y=[(\boldsymbol{R}^{\delta})^{\mathrm{H}},0](\boldsymbol{Q}^{\delta})^{\mathrm{H}}x^{\delta}=[(\boldsymbol{R}^{\delta})^{\mathrm{H}},0]\begin{bmatrix}t^{\delta}\\ z^{\delta}\end{bmatrix}=(\boldsymbol{R}^{\delta})^{\mathrm{H}}t^{\delta}. \tag{5}$$

由(4)得
$$\|\boldsymbol{A}^{\delta}y-b^{\delta}\|_2=\|\boldsymbol{A}_1^{\delta}t^{\delta}+\boldsymbol{A}_2^{\delta}z^{\delta}-b^{\delta}\|_2.$$

由(5)得
$$\boldsymbol{B}^{\delta}y=d^{\delta}\Leftrightarrow(\boldsymbol{R}^{\delta})^{\mathrm{H}}t^{\delta}=d^{\delta}\Leftrightarrow\boldsymbol{R}^{\mathrm{H}}t=d.$$

可由 $\boldsymbol{R}^{\mathrm{H}}t=d$ 解出 t,从而
$$\min\|\boldsymbol{A}^{\delta}y-b^{\delta}\|_2=\min_{(R^{\delta})^{H}t^{\delta}=d^{\delta}}\|\boldsymbol{A}_1^{\delta}t^{\delta}+\boldsymbol{A}_2^{\delta}z^{\delta}-b^{\delta}\|_2=\min_{R^{H}t=d}\|\boldsymbol{A}_1^{\delta}t^{\delta}+\boldsymbol{A}_2^{\delta}z^{\delta}-b^{\delta}\|_2=$$
$$\min_{z^{\delta}}\|\boldsymbol{A}_2^{\delta}z^{\delta}-(b^{\delta}-\boldsymbol{A}_1^{\delta}t^{\delta})\|_2.$$

由[3]知 $\min\limits_{z^{\delta}}\|\boldsymbol{A}_2^{\delta}z^{\delta}-(b^{\delta}-\boldsymbol{A}_1^{\delta}t^{\delta})\|_2$ 的解为
$$z^{\delta}=(\boldsymbol{A}_2^{\delta})^{+}(b^{\delta}-\boldsymbol{A}_1^{\delta}t^{\delta})+(\boldsymbol{I}_{2n}-(\boldsymbol{A}_2^{\delta})^{+}\boldsymbol{A}_2^{\delta})\boldsymbol{K}.$$
其中 $\boldsymbol{K}$ 为任意 $2n\times2$ 矩阵.

取 $\boldsymbol{K}=\left[\begin{bmatrix}h_{st} & l_{st}\\ -\bar{l}_{st} & \bar{x}_{st}\end{bmatrix}\right]\in\mathbf{C}^{2n\times2}$,令 $c_{st}=h_{st}+l_{st}j$,

则 $\boldsymbol{C}=(c_{st})\in Q^{n\times1}$ 且 $\boldsymbol{C}^{\delta}=\boldsymbol{K}$.

即有 $$z^{\delta}=[\boldsymbol{A}_2^{+}(b-\boldsymbol{A}_1t)+(\boldsymbol{I}_n-\boldsymbol{A}_2^{+}\boldsymbol{A}_2)\boldsymbol{C}]^{\delta},$$

从而 $$z=\boldsymbol{A}_2^{+}(b-\boldsymbol{A}_1t)+(\boldsymbol{I}_n-\boldsymbol{A}_2^{+}\boldsymbol{A}_2))\boldsymbol{C},$$

所以(1)的解为
$$x=\boldsymbol{Q}\begin{bmatrix}t\\ z\end{bmatrix}.$$

最后,我们给出利用 OR 分解求解四元数矩阵的等式约束最小二乘问题的算法.

(1) 利用定理 2.2 对 $\boldsymbol{B}^{\mathrm{H}}$ 进行 OR 分解;

(2) 根据(5)将 $\boldsymbol{B}^{\delta}y=d^{\delta}$ 转化为上三角方程组 $R^{\mathrm{H}}t=d$ 求解 t;

(3) 令 $\boldsymbol{A}^{\delta}\boldsymbol{Q}^{\delta}=[\boldsymbol{A}_1^{\delta},\boldsymbol{A}_2^{\delta}]$,其中 $\boldsymbol{A}_1^{\delta}$ 为 $2p$ 列,求 $\boldsymbol{A}_2^{+}$ 并带入 $z=\boldsymbol{A}_2^{+}(b-\boldsymbol{A}_1t)+(\boldsymbol{I}_n-\boldsymbol{A}_2^{+}\boldsymbol{A}_2)\boldsymbol{C}$ 中,其中 $\boldsymbol{C}$ 为任意 $n\times1$ 四元数矩阵;

(4) 求出(1)的解为
$$x=\boldsymbol{Q}\begin{bmatrix}t\\ z\end{bmatrix}.$$

参考文献:

[1] 姜同松.矩阵的表示理论及其在数值计算中的应用[D].上海:华东师范大学数学系,2002.
[2] Tongsong Jiang, Musheng Wei. Equality constrained least squares problem over quaternion field[J]. Applied Mathematics Letters, 2003, (16):883~888.
[3] Janovska, G Opfer. Givens' transformation applied to quaternion valued vectors[J]. BIT Numerical Mathematics, 2003, (43):991~1 002.
[4] G W Stewart, Jiguang Sun. Matrix perturbation theory[M]. Academic Press, Inc, 1990.
[5] 连德忠.四元数体上的范数理论和应用[J].闽西职业大学学报,2003,(2):90~93.
[6] 姜同松.四元数的一种新的代数结构[J].力学学报,2002,(1):116~121.
[7] 李文亮.四元数矩阵[M].哈尔滨:国防科技大学出版社,2002.
[8] 李 莹,赵建立.四元数矩阵的 Rayleigh-Ritz 定理的证明[J].内蒙古大学学报,2006(1):5~8.

(编辑:李晓红)